Title

Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity.

Journal

Journal of medicinal chemistry, 52(3)

ISSN

0022-2623

Authors

Kerr, Iain D
Lee, Ji H
Pandey, Kailash C
et al.

Publication Date

2009-02-01

DOI

10.1021/jm8013663

Peer reviewed

Structures of Falcipain-2 and Falcipain-3 Bound to Small Molecule Inhibitors: Implications for Substrate Specificity‡

Iain D. Kerr,[†] Ji H. Lee,[†,#] Kailash C. Pandey,[§] Amanda Harrison,[§] Mohammed Sajid,[||] Philip J. Rosenthal,[§] and Linda S. Brinen*,[†]

Department of Cellular and Molecular Pharmacology and Department of Pathology, University of California, San Francisco, California 94158, and the Department of Medicine, San Francisco General Hospital, University of California, San Francisco, California 94143

Received October 28, 2008

Falcipain-2 and falcipain-3 are critical hemoglobinases of *Plasmodium falciparum*, the most virulent human malaria parasite. We have determined the 2.9 Å crystal structure of falcipain-2 in complex with the epoxysuccinate E64 and the 2.5 Å crystal structure of falcipain-3 in complex with the aldehyde leupeptin. These complexes represent the first crystal structures of plasmodial cysteine proteases with small molecule inhibitors and the first reported crystal structure of falcipain-3. Our structural analyses indicate that the relative shape and flexibility of the S2 pocket are affected by a number of discrete amino acid substitutions. The cumulative effect of subtle differences, including those at "gatekeeper" positions, may explain the observed kinetic differences between these two closely related enzymes.

Introduction

Approximately 300−500 million cases of malaria occur each year, leading to over 1 million deaths, nearly all from *Plasmodium falciparum*.[1] Widespread drug resistance increasingly limits the effectiveness of available therapies. New targets are required for the development of novel classes of antimalarial drugs.[2] Trophozoites of *P. falciparum* hydrolyze hemoglobin in an acidic food vacuole to generate free amino acids essential for parasite survival.[3,4] This process is blocked by the cysteine protease inhibitors E64 and leupeptin, as evidenced by the accumulation of undigested hemoglobin in the vacuole.[5]

The *P. falciparum* genome contains 33 open reading frames predicted to encode cysteine proteases, including 4 cathepsin L-like papain-family proteases collectively known as falcipains. Falcipain-1 (FP1[a]) has an uncertain role; although the protease is implicated in erythrocyte invasion, gene disruption has no effect on the development of erythrocytic parasites.[6,7] Disruption of falcipain-2 (FP2) leads to a marked reduction in hemoglobin hydrolysis by trophozoites, although this phenotype is transient;[8] disruption of the nearly identical falcipain-2′ (FP2′) has no apparent phenotype.[9] Disruption of falcipain-3 (FP3) has so far proved unsuccessful, and this result, taken together with the ready replacement of the gene by a plasmid-encoded copy, suggests an indispensable function for this protease.[9]

FP2 and FP3 share 66.7% sequence identity. This high level of similarity and a preference for substrates with a hydrophobic residue at the P2 position identifies these enzymes as cathepsin L-like. FP2 and FP3 contribute more or less equally to the digestion of hemoglobin in the food vacuole, with FP2 expressed earlier in the parasite life cycle.[10] Comparisons of FP2 and FP3 in the presence of smaller ligands show that FP3 is far less active against peptide substrates[10,11] and is usually less amenable to inhibition by peptidyl-based small molecules. We report the 2.9 Å crystal structure of FP2 in complex with E64 and the 2.5 Å crystal structure of FP3 in complex with leupeptin, providing detailed analyses of the active sites in the presence of small molecule inhibitors.

Results

Overall Structure. Cocrystallized FP2−E64 consists of residues 1−241 of the mature enzyme. The $C222_1$ asymmetric unit contains four molecules. Superimposition of the two most complete monomers (A and B) matches 235 α-carbons with root-mean-square distances (rmsd) of 0.56 Å. Residues 1−16 and 185−198 are located in insertions novel to plasmodial cysteine proteases and are more flexible than the rest of the structure. These two regions are only partially built in some monomers, and their exclusion matches 210 α-carbons with an rmsd of 0.45 Å.

FP3−leupeptin crystals belong to spacegroup *R*32 with two molecules of the mature enzyme, comprising residues 8−250, in the asymmetric unit. The two copies of the complex superimpose with an rmsd of 0.51 Å over 232 α-carbons. FP3 contains insertions equivalent to residues 1−16 and 185−198, and their exclusion matches 210 α-carbons with an rmsd of 0.45 Å. Residues 195−203 exhibit a higher degree of flexibility when compared to the surrounding structure. These residues are not well resolved in monomer B of FP3−leupeptin and are therefore not included in the model. Unless otherwise indicated, and for the sake of simplicity, our structural analyses involve monomer B from FP2 and monomer A from FP3, the best resolved monomer from each complex.

The Dali server[12] identifies cruzain, the major cysteine protease of *Trypanosoma cruzi*, as the parasitic homologue with the closest structural identity to both FP2 and FP3. All three share the common and well-characterized structural features of the two-domain papain-like fold. The core structure of the

‡ Coordinates for the falcipain-2 E64 and falcipain-3 leupeptin complexes have been deposited in the Protein Data Bank under accession numbers 3BPF and 3BPM, respectively.

* To whom correspondence should be addressed. Phone: 415-514-3426. Fax: +1 415-502-8193. E-mail: brinen@cmp.ucsf.edu.

† Department of Cellular and Molecular Pharmacology, University of California, San Francisco.

Current address: CrystalGenomics, Inc., 6F, 2nd Building, Asan Life Sciences Institute, Seoul 138-878, South Korea.

§ San Francisco General Hospital, University of California, San Francisco.

|| Department of Pathology, University of California, San Francisco.

[a] Abbreviations: FP1, falcipain-1; FP2, falcipain-2; FP2′, falcipain-2′; FP3, falcipain-3; LLG, log likelihood gain; TLS, translation−libration−screw; rmsd, root-mean-square distance.

10.1021/jm8013663 CCC: $40.75
Published on Web 01/07/2009

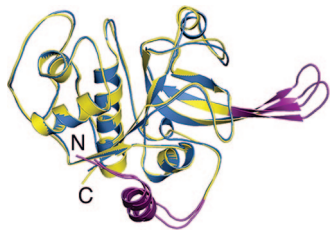

Figure 1. Structures of FP2 and FP3: superimposition of FP2−E64 and FP3−leupeptin with FP2 colored blue and FP3 colored yellow. Insertions are colored purple, and the N and C termini are labeled. All structure figures were prepared in PyMOL.[36]

domains is essentially identical to that found in all previously determined FP2 structures,[13−15] with minor variations in the insertions and loop regions. FP2 and FP3 share a high degree of structural similarity, and superimposition, using the DaliLite server,[12] matches 236 α-carbons with an rmsd of 0.8Å and a *Z*-score of 39.6 (Figure 1). The active site in each enzyme forms an extended, accessible cleft and is detailed in later sections.

FP2 and FP3 Share Two Unique, Structurally Equivalent Insertions. FP2 and FP3 share two unique insertions that distinguish plasmodial cysteine proteases from all other structurally characterized papain-family enzymes (Figure S1 in Supporting Information). The first insertion is located at the N-terminus of the mature enzyme and comprises residues 1−16 in FP2 and residues 1−25 in FP3. Residues 12−17 (monomer C) and 12−19 (monomer D) in FP2 lacked sufficient electron density to be included in the final models. Despite poor sequence identity between FP2 and FP3 in this region, the loop−helix−loop topology of the insertion (Figure 1) is well conserved and similar to that described previously.[13,15] The N-terminal insertion was previously shown to mediate folding of mature FP2.[4,16] Given the conserved structure of the insertion, it seems likely that it plays a similar role in FP3. Analogous to a similar interaction in FP2,[15] Tyr12 in FP3 provides a hydrogen-bonding anchor to Glu147, helping to stabilize enzyme structure in this region. A second possible anchor to Glu147, Lys20, is too distant to participate in a hydrogen bonding interaction (4.1 Å in monomer A, 4.4 Å in monomer B) and instead interacts with a sulfate anion that is an artifact of crystallization.

The second insertion in FP2 and FP3 forms a 14-residue β-hairpin at the C-terminus. This portion of FP2 was implicated in hemoglobin binding,[17] and the secondary structure in this region is well conserved between FP2 and FP3 (Figure 1). As previously noted for FP2,[13,15] this insertion has a tendency to be flexible, and this is the case in our FP2−E64 and FP3−leupeptin complexes. Accordingly, residues 189−193 (monomer A) and 182−198 (monomer D) in FP2 and 195−203 (monomer B) in FP3 lacked sufficient electron density to be included in the final models.

Active Sites of FP2 and FP3. The active sites of both enzymes are located in a cleft between the structurally distinct domains of the papain-like fold. Leupeptin and E64 interact with residues in the S1, S2, and S3 subsites of FP2 and FP3, corresponding to the P1, P2, and P3 positions of the ligands. The conserved catalytic residues of FP2 and FP3 (Gln36/45, Cys42/51, His174/183, Asn204/213, respectively) are similarly oriented with respect to the cocrystallized inhibitors. The active site cysteine forms a covalent, irreversible hemithioketal with the E64 epoxy carbon in the FP2−E64 complex and a covalent, reversible hemithioacetal with the asymmetric carbonyl carbon of leupeptin in the FP3−leupeptin complex (Figure 2).

As is commonly observed among inhibitors of papain-family enzymes, the carboxyl group of E64 and carbonyl group of

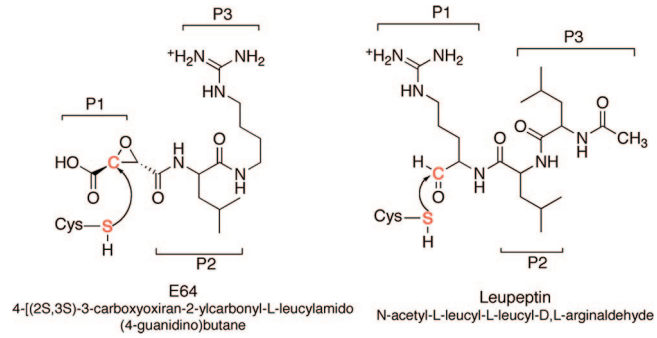

Figure 2. Chemical structures of E64 and leupeptin. The positions that occupy the S1, S2, and S3 subsites (P1, P2, and P3, respectively) are labeled. Enzyme and inhibitor groups involved during covalent bond formation are highlighted in red.

leupeptin occupy the oxyanion hole formed by the conserved catalytic residues (Figure S2 in Supporting Information). The (4-guanidino) butane moiety in E64 and the arginal moiety in leupeptin extend out into the solvent and are less ordered than the rest of the inhibitor molecules. Accordingly, these moieties are modeled at low occupancy in monomer B of both FP2 and FP3.

The peptidyl small molecule inhibitors are tethered to the main chains of FP2 and FP3 through a glycine residue that is highly conserved in the S3 subsite of clan CA cysteine proteases (Gly83 in FP2 and Gly92 in FP3). In each complex, this residue forms hydrogen bonds with the O and N atoms of the inhibitor backbone, similar to the pattern seen in β-sheet formation in protein secondary structure (parts A and B of Figure 3). In the FP2 active site, Gln36, Ser41 (monomer C only), Cys42, Asn81 (monomer A only), and His174 are involved in the formation of additional hydrogen bonds with E64[18] (Table S1 in Supporting Information). In the FP3−leupeptin complex, Gln45, Cys51, and Asn182 also act as hydrogen-bonding partners to the inhibitor (Table S1 in Supporting Information). A series of possible hydrophobic interactions are found between enzyme and inhibitor, involving the nonpolar regions of Gly40, Tyr78, Gly82, Leu84, Ser149, Leu172, Asn173, and Ala175 in FP2 and Tyr90, Gly91, Tyr93, Ile94, and Ser158 in FP3. Several hydrogen bonds are formed between FP3 and leupeptin via bridging water molecules; however, the interactions are not conserved in both copies of the complex and are likely not ligand dependent.

Discussion

The structures of FP2 and FP3 have been determined in complex with the small molecule inhibitors E64 and leupeptin, respectively. Both ligands display binding modes with their partners similar to those found in other papain-family enzymes, with these small molecule inhibitors largely targeting the S2 and S3 subsites. The P2 position in peptidyl ligands is the key determinant of substrate specificity in clan CA cysteine proteases.[19] FP2 and FP3 have a clear preference for substrates that contain a leucyl at this position (overall Leu > Phe > Val).[10,20] Both leupeptin and E64 contain a Leu at P2, and this residue slots comfortably into the active site to be involved in a number of hydrophobic interactions (Figure 3C).

FP2 and FP3 appear to perform similar functions in the trophozoite food vacuole, and this is reflected in the high degree of structural similarity shared by the two enzymes. Interestingly, FP3 has been shown to be much less efficient at processing peptide substrates and more difficult to inhibit with peptidyl-based small molecules.[10,11] The secondary structure that forms

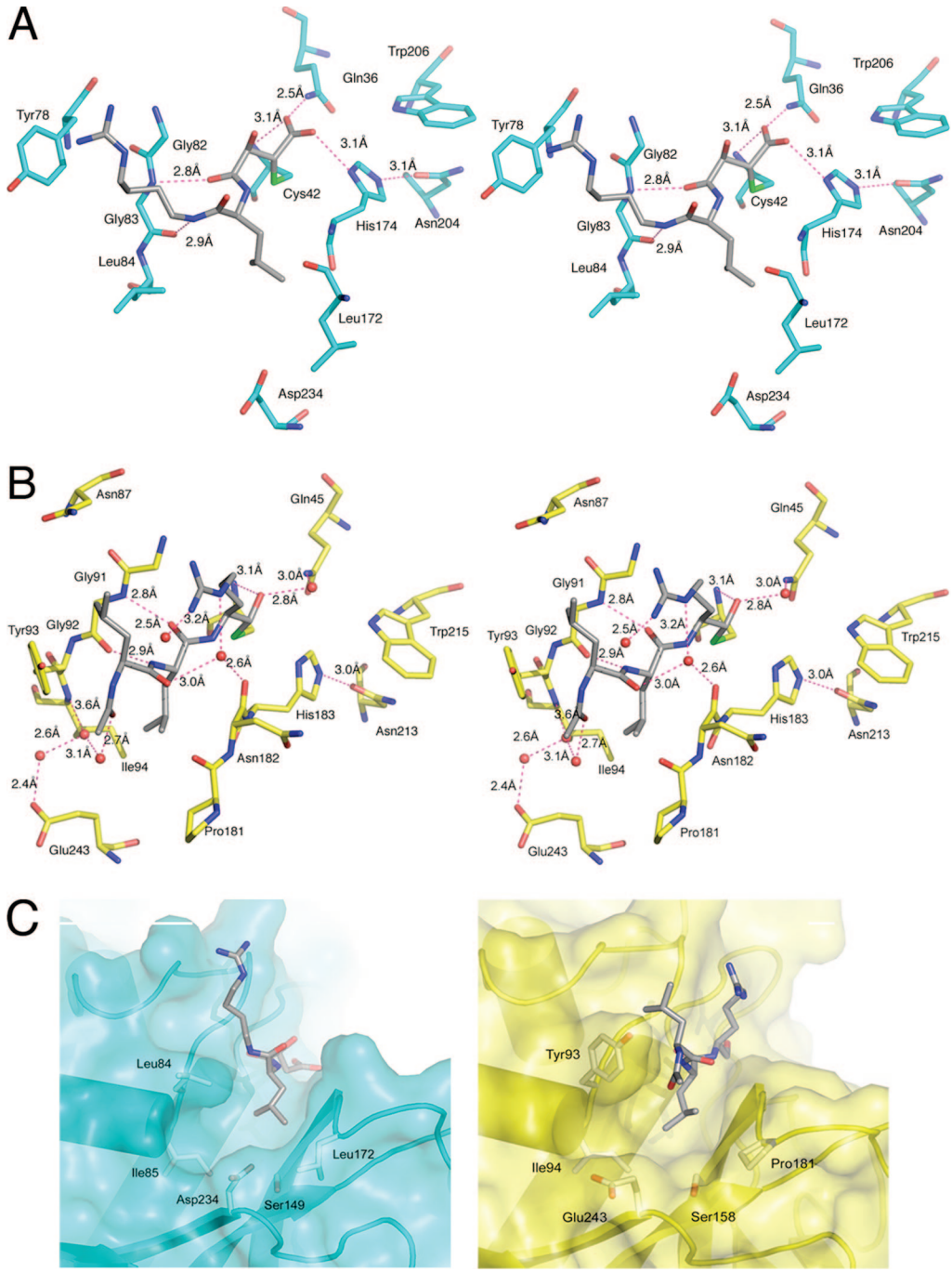

Figure 3. Active sites of FP2 and FP3. (A) FP2−E64 complex. Important residues in the active site are colored blue and labeled. E64 is least flexible in monomer A (shown here) and colored gray. Interactions with the enzyme are in pink. (B) FP3−leupeptin complex. Important residues in the active site are colored yellow and labeled. Leupeptin is colored gray, and interactions with the enzyme are in pink. (C) Surface representations of FP2 (left) and FP3 (right) highlighting the contour of the S2 subsite and important residues therein. Ligands are depicted as in (A) and (B).

the active site is highly conserved between the two enzymes, with no notable movements in the peptide backbone or loop regions. However, facile superimposition of FP2 and FP3 highlights two important substitutions in the S2 subsite; Asp234 vs Glu243 and Leu84 vs Tyr93, respectively. The amino acid at the Asp234/Glu243 position is nonconserved across the clan CA enzymes. This position lines the base of the S2 subsite and is known to be a key determinant of specificity. Both leupeptin and E64 have a leucyl at P2 and are therefore unable to form a charge−charge/hydrogen bonding interaction with an acidic residue at the bottom of the S2 subsite. In the FP2−E64 complex, Asp234 is rotated away from the inhibitor, as is seen in the complexes with chagasin (Leu at P2[14]) and cystatin (no P2 residue[15]). This residue is restricted from pointing out toward the solvent through either a hydrogen bonding interaction with the nearby Ser149 in the S2 subsite (monomer A) or interaction with Asn86 through a bridging water molecule (monomer B). Asp234 in monomers C and D possibly also form a stabilizing hydrogen bond with Ser149; however, at 3.7 and 3.6 Å, respectively, the distances are slightly too long for bond formation.

Glu243 in the FP3−leupeptin complex also points away from the inhibitor but, in contrast to the FP2−E64 complex, does point out toward the solvent. The conformation of the side chain is similar in both copies of the complex in the asymmetric unit and is stabilized in both monomers through the formation of a

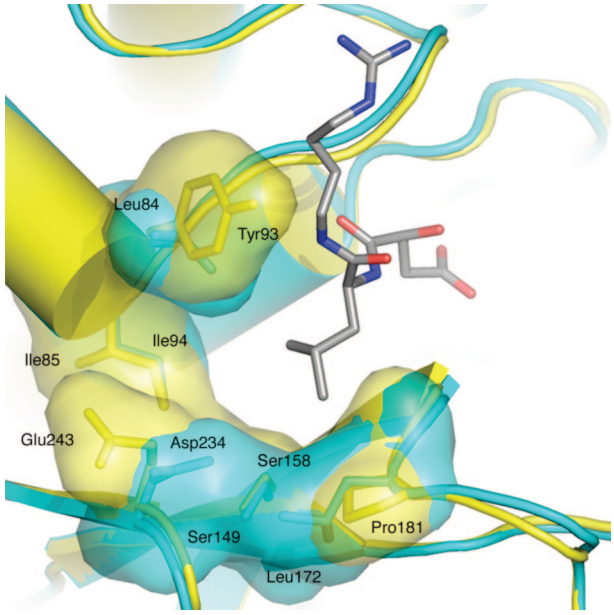

Figure 4. S2 subsites of FP2 and FP3: surface representation of the S2 subsites of FP2 and FP3. Colors are in accordance with Figure 3. E64 is depicted in stick form.

direct hydrogen bond to Tyr245. This part of the S2 subsite is more solvent exposed in monomer A, allowing additional stabilizing interactions through bridging water molecules with Ile94, Thr95, and the *N*-acetylcarbonyl of leupeptin. In comparison to Asp234 in FP2, the additional side chain carbon (and thus rotatable bond) in Glu243 of FP3 increases the size of the residue and the degree of side chain flexibility at this position. In the FP3−leupeptin structure, the outcome of a larger, more flexible residue at the bottom of the S2 subsite and the bulkier Leu84 to Tyr93 substitution above is to narrow the wall of the S2 subsite formed by Tyr93, Ile94, and Glu243 when compared with the same region in FP2 (Leu84, Ile85, and Asp234) (Figure 4).

A homology model of FP3[21] previously pointed to a role for Tyr93 in restricting parts of the S2 subsite. The substitution of Leu172 in FP2 to Pro181 in FP3 was also implicated and, in combination with Tyr93, suggested to be directly responsible for the narrowing of the S2 subsite. In the *Fasciola hepatica* cathepsin L1, these two positions are referred to as "gatekeepers" because they sit at the entrance to the S2 subsite.[22] Tyr93, by virtue of its size, may play a direct role in narrowing the S2 subsite. The role of Pro181, which has a substantially restrained set of peptide ϕ/Ψ combinations compared to the other amino acids, may be to reduce the conformational flexibility around the entrance to the S2 subsite. The combination of a bulky Tyr and a rigid Pro at "gatekeeper" positions serves to restrict the types of ligands that are able to access the entrance to the active site. As the major component of specificity is located at the S2 position of the enzyme, it is possible that amino acid changes of this nature could restrict P2 accessibility, thereby influencing catalytic efficiency and ligand recognition.

Overall, the complexes of FP2 and FP3 with E64 and leupeptin show a simpler mode of binding and inhibition compared to the previously determined crystal structures of FP2 with the macromolecular inhibitors chagasin and human cystatin.[14,15] The macromolecules adopt a tripartite mode of binding to FP2, burying a much larger surface area than is seen in either of our small molecule complexes. A wedge formed by three protease binding loops (BC, DE, and FG) allows chagasin to readily access both the prime and nonprime sites of the active site. Cystatin is only able to access the solvent exposed periphery of the nonprime sites and the majority of its binding to FP2 occurs at the prime end of the active site.

An interesting feature of the FP2−chagasin interaction is the highly mobile DE loop that, like E64, occupies the nonprime sites. In chagasin, the DE loop residues Leu64 and Leu65 correspond to the P3 and P2 positions of the ligand, respectively, and the peptide backbones occupy similar positions with respect to the conserved glycine residues in the S3 subsite (Figure S3 in Supporting Information). Furthermore, mutational analyses of ICP, a homologous macromolecular protease inhibitor from *Leishmania mexicana*,[23] highlight the importance of additional residues in the DE loop. The triple mutants Gly69Pro, Gly71Pro, Gly72Pro and Met67Asp, Val68Asp, Val70Asp abolished inhibitory activity.

Similar to E64 and leupeptin, both chagasin and cystatin inhibit FP2 and FP3 with nanomolar affinity. However, small molecule inhibitors are much more attractive as potential leads for the treatment of malaria. Unlike protein-based drugs, small molecule therapeutics are generally much more bioavailable, amenable to oral or topical administration, and vastly easier to redesign. The structures described in this report should assist ongoing efforts to develop falcipain inhibitors as new antimalarial drugs.

Experimental Procedures

Expression and Purification of the Mature FP2 Enzyme. FP2 was expressed in *E. coli* strain M15(pREP4) harboring the hexa-His-tagged pQ-35FP2 construct. Overexpression, refolding, and purification were carried out according to published protocols.[20]

The pH of refolded FP2 was adjusted to 5.8, DTT was added to a final concentration of 2.5 mM, and the solution was incubated at room temperature for 40 min. Following activation, hemoglobin and E64 (SIGMA) were added to a final concentration of 1 mM. Unbound hemoglobin and E64 were removed using a 10 mL Q-Sepharose column, and protein was eluted with a high salt buffer (20 mM Bis-Tris, pH 6.5, 0.5 M NaCl). The kinetic activity of purified FP2 was assayed as previously described[17] and the protein concentrated to approximately 7.4 mg/mL.

Expression and Purification of the Mature FP3 Enzyme. FP3 was expressed in *E. coli* strain M15(pREP4) harboring the hexa-His-tagged FP3-pQE-30 construct. Overexpression, refolding, and purification were carried out according to published protocols.[10]

The activity of FP3 was tested as for FP2 and was completely abolished by the addition of leupeptin (SIGMA) to a final concentration of 136 μM. Inhibited FP3 was purified using a 10 mL Q-sepharose column, and protein was eluted with high salt buffer (20 mM Bis-Tris, pH 6.5, 0.5 M NaCl). Fractions that contained FP3 were verified by SDS−PAGE, and buffer was exchanged with 20 mM Bis-Tris, pH 6.5, and concentrated to approximately 5 mg/mL.

Crystallization of FP2−E64 and FP3−Leupeptin Complexes. Our crystallization of the FP2−E64 complex was serendipitous, as initial trials were aimed at crystallizing FP2−E64 in complex with its natural substrate, hemoglobin. Crystals of what were believed to be the ternary complex were obtained using the sitting-drop vapor diffusion method[24] by equilibrating a mixture of 1 μL of protein−inhibitor complex (7.4 mg/ml) and 1 μL of reservoir solution (10% PEG 3350, 100 mM sodium nitrate, pH 6.0, 100 mM magnesium formate, 5% glycerol) against 500 μL of reservoir solution at 22 °C.

FP3−leupeptin crystals were obtained using the sitting-drop vapor diffusion method equilibrating 1 μL of protein−inhibitor complex (3 mg/mL) and 1 μL of reservoir solution (1.3 M ammonium sulfate, 50 mM sodium cacodylate, pH 6.5, 15 mM magnesium acetate) against 1 mL of reservoir solution at 22 °C

Data Collection and Structure Determination. Crystals were cryoprotected in mother liquor supplemented with 25% (FP2−E64)

Table 1. Data Collection and Refinement Statistics

Data collection	FP2−E64	FP3-leupeptin
space group	$C222_1$	$R32$
cell dimensions	a = 143.78 Å, b = 167.81 Å, c = 177.76 Å, $\alpha = \beta = \gamma = 90°$	$a = b$ = 154.57 Å, c = 129.01, $\beta = 120°$
resolution (Å)	2. 90 (3.10−2. 90)	2.50 (2.64−2.50)
Matthews coefficient (Å^3 Da^{-1})	4.94	2.65
solvent content (%)	75.1	53.6
R_{merge} [a] (%)	21.3 (83.5)	8.8 (49.1)
R_{pim} [b] (%)	8.0 (81.2)	2.3 (12.4)
redundancy	7.3 (7.5)	15.9 (16.3)
completeness (%)	100 (100)	100 (100)
$I/\sigma I$	10.8 (1.5)	27.0 (7.1)
Refinement		
R_{free} (%)	32.5	22.4
R_{factor} (%)	27.5	19.0
average B-factor (Å^2)	72	30
rmsd		
bonds (Å)	0.008	0.015
angles (deg)	1.04	1.54
Ramachandran plot[c]		
favored (%)	89.0	95.3
allowed (%)	98.1	99.8
outliers (%)	1.8	<0.01
PDB code	3BPF	3BPM

[a] $R_{merge} = \sum\sum[I(h)j - I(h)]/\sum\sum I(h)j$ where $I(h)$ is the measured diffraction intensity and the summation includes all observations. [b] $R_{pim} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i|Ii(hkl) - I(hkl)|\sum_{hkl}\sum_i Ii(hkl)$.[29] [c] As defined by Molprobity.[35]

and 15% ethylene glycol (FP3−leupeptin), respectively. Samples were then mounted in standard Hampton cryoloops, flash-cooled in liquid nitrogen, and loaded into a Stanford automated mounting system (SAM) sample cassette.[25] Data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-1, Menlo Park, CA, on an ADSC Q-315 detector.

Diffraction data were collected from a single FP2−E64 crystal as 0.5°, 35 s oscillations. FP2−E64 crystals belong to space group $C222_1$ (a = 143.78 Å, b = 167.81 Å, c = 177.76 Å). Diffraction data from a single FP3−leupeptin crystal were collected as 1°, 20 s oscillations. FP3 crystals belong to space group $R32$ ($a = b$ = 154.57 Å, c = 129.01 Å, $\beta = 120°$). Reflections were indexed and integrated using XDS[26] for FP2−E64 and MOSFLM for FP3−leupeptin.[27] Data were scaled and merged with XSCALE for FP2−E64[26] and SCALA for the FP3−leupeptin complex.[28] Because of the high redundancy of the data, we have included the precision indicating merging R-factor, R_{pim}[29] (calculated in SCALA), in Table 1 as a more accurate description of the precision of the averaged measurements. Structure factor amplitudes were calculated using TRUNCATE.[30]

The structure of FP2−E64 was determined by molecular replacement in PHASER[31] using the FP2 component of the FP2−cystatin complex (PDB 1YVB).[15] Four independent monomers were located in the asymmetric unit, yielding a solution with a log likelihood gain (LLG) of 4982 and translation function Z-score of 50.2. Following initial refinement in CNS,[32] $mF_o - DF_c$ SIGMAA-weighted electron density maps confirmed that hemoglobin was absent. However, the presence of E64 was clear and we were able to position the small molecule inhibitor in the active site of all four monomers. The model was refined over several rounds in CNS, interspersed with manual adjustments in COOT.[33] Refinement was concluded in REFMAC5 using TLS parametrization.[34] The model shows good stereochemistry, as assessed by MOLPROBITY,[35] and is refined to a final R_{free} and R-factor of 32.5% and 27.5%, respectively. Structure statistics (PDB 3BPF) are summarized in Table 1.

The structure of FP3 was determined by molecular replacement in PHASER using one molecule of the FP2−E64 complex. Two monomers were located in the asymmetric unit, yielding a solution with a log likelihood gain (LLG) of 3661 and translation function Z-score of 59.9. Refinement/rebuilding of the model and placement of ligands followed the same protocol as the FP2−E64 structure. The model shows good stereochemistry, as assessed by MOLPROBITY, and is refined to a final R_{free} and R-factor of 22.4% and 19.0%, respectively. Structure statistics (PDB 3BPM) are summarized in Table 1.

Acknowledgment. Portions of this work were supported by National Institutes of Health (Awards AI35707 and AI35800), the Sandler Foundation, and the Medicines for Malaria Venture. The authors thank Nestor O. Choncha, Thomas Stout, Max Nanao, Irimpan Mathews, Tzanko Duokov, Chris Waddling, Jennifer Legac, and Jiri Gut for assistance with experiments. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. P.J.R. is a Doris Duke Charitable Foundation Distinguished Clinical Scientist.

Supporting Information Available: Table of inhibitor−enzyme interactions, a structure-based sequence alignment of FP2 and FP3, E64 and leupeptin omit maps, and a comparison of small molecule inhibitors and chagasin. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Guinovart, C.; Navia, M. M.; Tanner, M.; Alonso, P. L. Malaria: burden of disease. *Curr. Mol. Med.* **2006**, *6*, 137–140.
(2) Gelb, M. H. Drug discovery for malaria: a very challenging and timely endeavor. *Curr. Opin. Chem. Biol.* **2007**, *11*, 440–445.
(3) Francis, S. E.; Sullivan, D. J., Jr.; Goldberg, D. E. Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. *Annu. Rev. Microbiol.* **1997**, *51*, 97–123.
(4) Rosenthal, P. J. Hydrolysis of erythrocyte proteins by proteases of malaria parasites. *Curr. Opin. Hematol.* **2002**, *9*, 140–145.
(5) Rosenthal, P. J.; McKerrow, J. H.; Aikawa, M.; Nagasawa, H.; Leech, J. H. A malarial cysteine proteinase is necessary for hemoglobin degradation by *Plasmodium falciparum*. *J. Clin. Invest.* **1988**, *82*, 1560–1566.
(6) Greenbaum, D. C.; Baruch, A.; Grainger, M.; Bozdech, Z.; Medzihradszky, K. F.; Engel, J.; DeRisi, J.; Holder, A. A.; Bogyo, M. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* **2002**, *298*, 2002–2006.
(7) Rosenthal, P. J.; Nelson, R. G. Isolation and characterization of a cysteine proteinase gene of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **1992**, *51*, 143–152.
(8) Sijwali, P. S.; Rosenthal, P. J. Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4384–4389.
(9) Sijwali, P. S.; Koo, J.; Singh, N.; Rosenthal, P. J. Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **2006**, *150*, 96–106.
(10) Sijwali, P. S.; Shenai, B. R.; Gut, J.; Singh, A.; Rosenthal, P. J. Expression and characterization of the *Plasmodium falciparum* haemoglobinase falcipain-3. *Biochem. J.* **2001**, *360*, 481–489.
(11) Ramjee, M. K.; Flinn, N. S.; Pemberton, T. P.; Quibell, M.; Wang, Y.; Watts, J. P. Substrate mapping and inhibitor profiling of falcipain-2, falcipain-3 and berghepain-2: implications for peptidase anti-malarial drug discovery. *Biochem. J.* **2006**, *399*, 47–57.
(12) Holm, L.; Sander, C. Mapping the protein universe. *Science* **1996**, *273*, 595–603.
(13) Hogg, T.; Nagarajan, K.; Herzberg, S.; Chen, L.; Shen, X.; Jiang, H.; Wecke, M.; Blohmke, C.; Hilgenfeld, R.; Schmidt, C. L. Structural and functional characterization of falcipain-2, a hemoglobinase from the malarial parasite *Plasmodium falciparum*. *J Biol Chem* **2006**, *281*, 25425–25437.
(14) Wang, S. X.; Pandey, K. C.; Scharfstein, J.; Whisstock, J.; Huang, R. K.; Jacobelli, J.; Fletterick, R. J.; Rosenthal, P. J.; Abrahamson,

M.; Brinen, L. S.; Rossi, A.; Sali, A.; McKerrow, J. H. The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. *Structure* **2007**, *15*, 535–543.

(15) Wang, S. X.; Pandey, K. C.; Somoza, J. R.; Sijwali, P. S.; Kortemme, T.; Brinen, L. S.; Fletterick, R. J.; Rosenthal, P. J.; McKerrow, J. H. Structural basis for unique mechanisms of folding and hemoglobin binding by a malarial protease. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11503–11508.

(16) Sijwali, P. S.; Shenai, B. R.; Rosenthal, P. J. Folding of the *Plasmodium falciparum* cysteine protease falcipain-2 is mediated by a chaperone-like peptide and not the prodomain. *J. Biol. Chem.* **2002**, *277*, 14910–14915.

(17) Pandey, K. C.; Wang, S. X.; Sijwali, P. S.; Lau, A. L.; McKerrow, J. H.; Rosenthal, P. J. The *Plasmodium falciparum* cysteine protease falcipain-2 captures its substrate, hemoglobin, via a unique motif. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9138–9143.

(18) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein−ligand interactions. *Protein Eng.* **1995**, *8*, 127–134.

(19) Storer, A. C.; Menard, R. Recent insights into cysteine protease specificity: lessons for drug design. *Perpect. Drug Discovery Des.* **1996**, *6*, 33–46.

(20) Shenai, B. R.; Sijwali, P. S.; Singh, A.; Rosenthal, P. J. Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. *J Biol Chem* **2000**, *275*, 29000–29010.

(21) Sabnis, Y. A.; Desai, P. V.; Rosenthal, P. J.; Avery, M. A. Probing the structure of falcipain-3, a cysteine protease from *Plasmodium falciparum*: comparative protein modeling and docking studies. *Protein Sci.* **2003**, *12*, 501–509.

(22) Stack, C. M.; Caffrey, C. R.; Donnelly, S. M.; Seshaadri, A.; Lowther, J.; Tort, J. F.; Collins, P. R.; Robinson, M. W.; Xu, W.; McKerrow, J. H.; Craik, C. S.; Geiger, S. R.; Marion, R.; Brinen, L. S.; Dalton, J. P. Structural and functional relationships in the virulence-associated cathepsin L proteases of the parasitic liver fluke, *Fasciola hepatica*. *J. Biol. Chem.* **2008**, *283*, 9896–9908.

(23) Smith, B. O.; Picken, N. C.; Westrop, G. D.; Bromek, K.; Mottram, J. C.; Coombs, G. H. The structure of *Leishmania mexicana* ICP provides evidence for convergent evolution of cysteine peptidase inhibitors. *J. Biol. Chem.* **2006**, *281*, 5821–5828.

(24) McPherson, A. *Preparation and Analysis of Protein Crystals*; Wiley: Chichester, U.K., 1982.

(25) Cohen, A. E.; Ellis, P. J.; Deacon, A. M.; Miller, M. D.; Phizackerley, R. P. An automated system to mount cryo-cooled protein crystals on a synchrotron beamline, using compact sample cassettes and a small-scale robot. *J. Appl. Crystallogr.* **2002**, *35*, 720–726.

(26) Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **1993**, *26*, 795–800.

(27) Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EAMCB Newsletter on protein Crystallography* **1992**, 26.

(28) Evans, P. R. SCALA. *Jt. CCP4 ESF-EAMCB Newsletter Protein Crystallogr.* **1997**, *33*, 22–24.

(29) Weiss, M. S.; Hilgenfeld, R. On the use of the merging *R* factor as a quality indicator for X-ray data. *J. Appl. Crystallogr.* **1997**, *30*, 203–205.

(30) French, S.; Wilson, K. On the treatment of negative intensity observations. *Acta Crystallogr. A* **1978**, *34*, 517–525.

(31) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(32) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905–921.

(33) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(34) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240–255.

(35) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **2007**, *35*, W375−W383.

(36) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: San Carlos, CA, 2002.

JM8013663